

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2002, 08:24:57 ; Search time 76.45 Seconds  
(without alignments)  
174,239 Million cell updates/sec

Title: US-09-351-778a-11

Sequence: 1 MTSSTLAPTDYRNTATGCL.....LICCLKRRRAPSLLLOYD 77

Scoring table: OLIGO  
Gapop 60.0 , Capext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	83.1	101	12	Q91023
2	48	62.3	101	12	Q91256
3	18	2.4	94	12	Q12392
4	7	9.1	83	12	Q88422
5	7	9.1	87	10	Q94H62
6	7	9.1	102	11	Q9D126
7	7	9.1	160	16	Q96268
8	7	9.1	160	16	Q92XL7
9	7	9.1	197	16	Q9JSG7
10	7	9.1	207	4	Q96DJ9
11	7	9.1	228	2	Q9R8X8
12	7	9.1	245	8	Q9T3Y9
13	7	9.1	249	12	Q92466
14	7	9.1	309	5	Q95WFS
15	7	9.1	316	10	Q9C926
16	7	9.1	387	16	Q99X15

17	7	9.1	391	16	Q97GC3	Q97GC3 clostridium
18	7	9.1	392	10	Q9AX94	Q9AX94 oryza sativ
19	7	9.1	404	5	Q9VIM7	Q9VIM7 drosophila
20	7	9.1	476	10	Q9LIP4	Q9LIP4 oryza sativ
21	7	9.1	490	10	Q9FSS8	Q9FSS8 oryza sativ
22	7	9.1	536	5	Q45994	Q45994 caenorhabd
23	7	9.1	635	5	Q9VJ77	Q9VJ77 drosophila
24	7	9.1	649	5	Q9NKP0	Q9NKP0 drosophila
25	7	9.1	665	2	Q92494	Q92494 bacteroides
26	7	9.1	1175	16	Q91356	Q91356 pseudomonas
27	7	9.1	1401	11	Q9JKD4	Q9JKD4 mus musculu
28	7	9.1	1847	5	Q9NKN5	Q9NKN5 leishmania
29	7	9.1	2499	5	Q9BN90	Q9BN90 homo sapien
30	6	7.8	26	4	Q9BUB3	Q9BUB3 homo sapien
31	6	7.8	53	8	Q78735	Q78735 flammulina
32	6	7.8	53	13	Q9VH36	Q9VH36 coturnix co
33	6	7.8	64	2	Q9P347	Q9P347 streptomyce
34	6	7.8	69	13	Q9DEC6	Q9DEC6 gallus galli
35	6	7.8	73	13	Q87603	Q87603 chimpanzee
36	6	7.8	76	4	Q96EN4	Q96EN4 homo sapien
37	6	7.8	77	12	Q64888	Q64888 avian adeno
38	6	7.8	79	12	Q9PX71	Q9PX71 human herpe
39	6	7.8	83	16	Q9PEU4	Q9PEU4 xylella fas
40	6	7.8	89	10	Q94T51	Q94T51 oryza sativ
41	6	7.8	89	16	Q9A887	Q9A887 caulobacter
42	6	7.8	91	4	Q9H4V4	Q9H4V4 homo sapien
43	6	7.8	92	2	Q93SB6	Q93SB6 frankia sp.
44	6	7.8	94	16	Q9PCB2	Q9PCB2 xylella fas
45	6	7.8	94	16	Q9PCAS	Q9PCAS xylella fas

## ALIGNMENTS

RESULT 1  
ID Q91023 PRELIMINARY: PRT: 101 AA.  
AC Q91023;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE 11.6K PROTEIN.  
OS Human adenovirus type 2.  
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
CN NCBI\_TaxId=10515;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=105;  
RA Borchering F., Ping-Akerblom P.;  
RT "Adenoviruses of subgenus C with different organ tropism."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ293915; CAC67721.1; -  
SQ SEQUENCE 101 AA; 11662 MW; 914F50AC2F8B284F CRC64;

Query Match 83.1%; Score 64; DB 12; Length 101;  
Best Local Similarity 100.0%; Pred. No. 4.8e-60;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 APTTDRNTATGTSALNLPQVAFVNDASLDMMFSTALMFVCLIMLICKLR 66  
DB 7 APTTDRNTATGTSALNLPQVAFVNDASLDMMFSTALMFVCLIMLICKLR 66  
QY 67 ARPP 70  
DB 67 ARPP 70  
RESULT 2  
ID Q91256 PRELIMINARY: PRT: 101 AA.  
AC Q91256;  
DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE 11.6K PROTEIN.  
 OS Human adenovirus type 2.  
 CC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 ON NCBI\_TaxID=10515;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PREI;  
 RA Borchering F., Ping-Akerblom P.;  
 RT "Adenoviruses of subgenus c with different organ tropism";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: A2793913; CAC67704.1; -;  
 SO SEQUENCE 101 AA; 11704 MW; E13857DC5891B85B CRC64;

Query Match 62.3%; Score 48; DB 12; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-43;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 ALNPOVHAENVDMASLDMMFSLMFCVLIIMVLCCLKRRRAPP 70  
 DB 23 ALNPOVHAENVDMASLDMMFSLMFCVLIIMVLCCLKRRRAPP 70

RESULT 3  
 ID 012392 PRELIMINARY; PRT: 94 AA.  
 AC 012392;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
 DT 01-JAN-1998 (Tremblrel. 05, Last annotation update)  
 DE 11.6K PROTEIN.  
 GN AD1/E3-11.6K.  
 OS Human adenovirus type 1.  
 CC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 ON NCBI\_TaxID=10533;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HANOVER / ADRIAN;  
 RA Reichmann H., Scharschmidt E., Geisler B., Hausmann J., Ortmann D.,  
 RA Bauer U., Flunker G., Seidel W.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y11257; CAA72127.1; -;  
 DR EMBL: Y11032; CAA71916.1; -;  
 SO SEQUENCE 94 AA; 10674 MW; D11485AFE771862 CRC64;

Query Match 23.4%; Score 18; DB 12; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-11;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 MMFSLMFCVLIIMVLCCLKRRRAPP 58  
 DB 35 MMFSLMFCVLIIMVLCCLKRRRAPP 52

RESULT 4  
 ID 088422 PRELIMINARY; PRT: 83 AA.  
 AC 088422;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE SPV1-C74, COMPLETE GENOME.  
 OS Spiroplasma virus.  
 CC Viruses; unclassified viruses.  
 ON NCBI\_TaxID=12338;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPV1-C74;  
 RA Bebear C.M., Aulio P., Bove J., Renaudin J.;  
 RT "Spiroplasma citri virus spv1. Characterization of viral sequences

RT present in the spiroplasma host chromosome."  
 RL Curr. Microbiol. 32:1-7(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPV1-C74;  
 RA Renaudin J.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U28974; AAA85015.1; -;  
 SO SEQUENCE 83 AA; 9398 MW; D4969373B02BFFA8 CRC64;

Query Match 9.1%; Score 7; DB 12; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 9.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 TGLTSAL 24  
 DB 30 TGLTSAL 36

RESULT 5  
 ID 094H62 PRELIMINARY; PRT: 87 AA.  
 AC 094H62;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE HYPOTHETICAL 9.9 KDA PROTEIN.  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Ehrhartoideae; Oryzaceae; Oryza.  
 ON NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPONBARE;  
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,  
 RA Brenner M., Burgess S., Hance M., Shwartsbeyn M., Tsitrin T.,  
 RA Riggs F., Hsiao J., Zisman V., Blunt S., Pal G., Vanaken S.E.,  
 RA Utechtack T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,  
 RA White O., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSNBNB0057P11 genomic sequence."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC084767; AAK72273.1; -;  
 KW Hypothetical protein.  
 SO SEQUENCE 87 AA; 9941 MW; 092E2A6E9D0A42CE CRC64;

Query Match 9.1%; Score 7; DB 10; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 9.8;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 LKRRAR 68  
 DB 48 LKRRAR 54

RESULT 6  
 ID 09D126 PRELIMINARY; PRT: 102 AA.  
 AC 09D126;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE A930014B1IRIK PROTEIN.  
 GN A930014B1IRIK.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=RETINA;

RX MEDLINE-21085660; PubMed-11217851;  
 RA Kawai J., Shingawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishii K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Pletschmann M., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schiraldi L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guadincioni S., Hill D., Hofmann M., Hune D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszewski A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK020853; BAB32229.1; -;  
 DR MGI: 1924349; A930014B11RLK.  
 SO SEQUENCE 102 AA; 11333 MW; 173C4E191A1477DB CRC64;

Query Match  
 Best Local Similarity 9.1%; Score 7; DB 11; Length 102;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 RPSLL 74  
 DB 86 RPSLL 92

RESULT 7  
 ID 098268 PRELIMINARY; PRT; 160 AA.  
 AC 098268;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DE 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE MR9079 PROTEIN.  
 GN MR9079.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MAFE303099;  
 RX MEDLINE-21082930; PubMed-11214968;  
 RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Matanabe A., Ideasa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL: AP003015; BAB54491.1; -;  
 KW Plasmid; Complete proteome.  
 SO SEQUENCE 160 AA; 17656 MW; 85A9C5A9233D9A50 CRC64;

Query Match  
 Best Local Similarity 9.1%; Score 7; DB 16; Length 160;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 TGSTIAP 8  
 DB 28 TGSTIAP 34

RESULT 8  
 ID 092XL7 PRELIMINARY; PRT; 160 AA.  
 AC 092XL7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE HYPOTHEICAL PROTEIN SMA2279.  
 GN SMA2279.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Bacteroidetes; Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1021;  
 RX MEDLINE-21396509; PubMed-11481432;  
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
 RA Barlett-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
 RA Gujal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,  
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
 RT "Nucleotide sequence and predicted functions of the entire  
 RT Sinorhizobium meliloti pSyma megaplasmid.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
 DR EMBL: AE007307; AAK65885.1; -;  
 KW Hypothetical protein; Plasmid; Complete proteome.  
 SO SEQUENCE 160 AA; 17966 MW; 9BABAEBE18A1BE34 CRC64;

Query Match  
 Best Local Similarity 9.1%; Score 7; DB 16; Length 160;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 TGSTIAP 8  
 DB 28 TGSTIAP 34

RESULT 9  
 ID 09JSG7 PRELIMINARY; PRT; 197 AA.  
 AC 09JSG7;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CT142 HYPOTHETICAL PROTEIN\_2.  
 GN CT142.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-J138;  
 RX MEDLINE-20330349; PubMed-10871362;  
 RA Shirai M., Hiraoka H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hatori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 RT from Japan and CW1029 from USA.";  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 DR EMBL: AP002545; BAA98469.1; -;  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR InterPro: IPR000531; TonB\_box.  
 DR PROSITE: PS00433; GRAM\_POS\_ANCHORING; UNKNOWN.1.  
 DR PROSITE: PS00430; TONB\_DEPENDENT\_REC.1; UNKNOWN.1.  
 SO SEQUENCE 197 AA; 22016 MW; EAA69A27851F790D CRC64;

Query Match  
 Best Local Similarity 9.1%; Score 7; DB 16; Length 197;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 ATGTTSA 23  
 |||||  
 DB 74 ATGTTSA 80

## RESULT 10

096DJ9 PRELIMINARY; PRT: 207 AA.

AC 096DJ9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CNA FLJ25310 F15, CLONE SYN00991.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-STNOVIAL MEMBRANE (KNEE);  
 RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,  
 RA Houta T., Hirooka S., Murakawa K., Takiguchi S., Kusano J.,  
 RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,  
 RA Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,  
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,  
 RA Kawakami B., Nagai K., Isegaki T., Sugano S.,  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK058039; BAB71635.1; -;  
 SO SEQUENCE 207 AA; 22532 MW; 8E8A17008E924F0 CRC64;

Query Match 9.1%; Score 7; DB 4; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 RPSLL 74  
 |||||  
 DB 68 RPSLL 74

## RESULT 11

09RBX8 PRELIMINARY; PRT: 228 AA.

AC 09RBX8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYPOHETICAL 24.1 KDA PROTEIN.  
 OS Pseudomonas indologera.  
 CC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Vogesella*.  
 OX NCBI\_TaxID=45465;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC19706;  
 RA van de Loo F.J., Keese P., Llewellyn D.;  
 RT "Structural and regulatory genes controlling indigoldine production in  
 RT *Vogesella indologera*: involvement of a peptide synthetase homolog";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC EMBL: AF088856; AAD54003.1; -;  
 DR InterPro: IPR001647; HTH\_Tetr.  
 DR Pfam: PF00440; tetr.1  
 DR PRINTS: PR00455; HTHTR.  
 KW DNA-binding; Hypothetical protein; Transcription regulation.  
 SO SEQUENCE 228 AA; 24140 MW; A928D14F404869B CRC64;

Query Match 9.1%; Score 7; DB 2; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 RRRARP 70  
 |||||  
 DB 23 RRRARP 29

## RESULT 12

09T3Y9 PRELIMINARY; PRT: 245 AA.

AC 09T3Y9;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 27.2 KDA PROTEIN.  
 OS Nephroselmis olivacea.  
 CC Chloroplast.  
 CC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;  
 CC Chlorodendreales; Chlorodendraceae; Nephroselmis.  
 OX NCBI\_TaxID=31312;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-9938694; PubMed-10468594;  
 RX Tumei M., Ots C., Lemieux C.;  
 RA "The complete chloroplast DNA sequence of the green alga *Nephroselmis*  
 RA *olivacea*. Insights into the architecture of ancestral chloroplast  
 RT genomes";  
 RT Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Tumei M., Ots C., Lemieux C.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF137379; AAD54918.1; -;  
 DR EMBL: AF137379; AAD54871.1; -;  
 KW Hypothetical protein; Chloroplast.  
 SO SEQUENCE 245 AA; 27217 MW; 74FDF5FD5F229FF7 CRC64;

Query Match 9.1%; Score 7; DB 8; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 TGLTSAL 24  
 |||||  
 DB 204 TGLTSAL 210

## RESULT 13

092466 PRELIMINARY; PRT: 249 AA.

AC 092466;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ACNMPV ORF106.  
 GN ORF.90.  
 OS Bombyx mori nuclear polyhedrosis virus (BmNPV).  
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 CC Nucleopolyhedrovirus.  
 OX NCBI\_TaxID=10436;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-T3;  
 RX MEDLINE-97329351; PubMed-9185864;  
 RA Kamita S.G., Maeda S.;  
 RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx  
 RT mori nuclear polyhedrosis virus and fine-mapping of a region involved  
 RT in host range expansion";  
 RL Gene 190:173-179(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-T3;  
 RX MEDLINE-99281911; PubMed-10355780;

```

RA Genl S., Matjima K., Maeda S.;
RT *Sequence analysis of the genome of Bombyx mori
RT nucleopolyhedrovirus."
RL J. Gen. Virol. 80:1323-1337(1999).
RN (13)
RP SEQUENCE FROM N.A.
RC STRAIN-T3;
RA Maeda S.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN (14)
RC SEQUENCE FROM N.A.
RC STRAIN-T3;
RA Genl S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: L33180; AAC63776.1; -
SO SEQUENCE 249 AA; 28921 MW; FC65F3EBBF298EF8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 249;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 ARPSLL 73
DB 194 ARPSLL 200

RESULT 14
O95WT5 PRELIMINARY; PRT; 309 AA.
ID O95WT5;
AC O95WT5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HOMEBOX HX.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OC NCBI_TaxID=7739;
OX (1)
RN SEQUENCE FROM N.A.
RP Kozmik Z., Vlcek C.;
RT "A novel homeobox gene in amphioxus."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF303216; AAL09323.1; -
SO SEQUENCE 309 AA; 34360 MW; B5ED048AF6E74FC CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 5; Length 309;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 69 PPSLLQ 75
DB 262 PPSLLQ 268

RESULT 15
O9C926 PRELIMINARY; PRT; 316 AA.
ID O9C926;
AC O9C926;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOPHYSICAL 34.7 KDA PROTEIN (ATGCG8600/F17014_7).
GN F17014.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid1; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-CV, COLUMBIA;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unselid M.,
RA Fattmann B., Valle G., Bloeker H., Perez-Alonso M., Oberaler B.,
RA Delseny M., Boutry M., Griuell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choisme N., Artiguenave F., Robert C., Brotlier P.,
RA Winkler P., Cuatrecasas L., Weissenbach J., Saurin W., Queller F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T., H., Nordstiek G.,
RA Reichelt J., Scharfe M., Schoen O., Bagnies M., Terol J., Clement J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemin D.,
RA Cooke R., Landie M., Berger-Liauro C., Purnelle B., Masu D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casachuberta E.,
RA Monfort A., Argitrou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Ruid S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Uterback T., Fujii C.Y., Shea T.P.,
RA Greasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltischer J., Sellers P., Gill J.E., Feldlyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Itoesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsunoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana."
RL Nature 408:820-822(2000).
RN (12)
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carinici P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamliya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN (13)
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carinici P., Dale J.M., Gibson H.A., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamliya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC012562; AAG51364.1; -
DR EMBL: AY056087; AAL06975.1; -
DR EMBL: AY045678; AAK74036.1; -
KW Hypothetical protein.
SO SEQUENCE 316 AA; 34732 MW; SB54FCF59A5B5B CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 316;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 SALNLPQ 28
DB 146 SALNLPQ 152

```

Search completed: June 21, 2002, 08:24:58

job time: 283 sec

---